Technical Note

1967-3

P. L. Fleck

FASTABUL

(A Fast Automatic STAtion BULletin Program)

6 January 1967

FASTABUL
(A Fast Automatic STAtion BULletin Program)

P. L. FLECK

Group 64

ABSTRACT

A real-time program for a digital computer is described which uses the sampled outputs of an array of 21 seismometers as input. These data are reduced and output as a typewritten list of the physical parameters of any "event" present in the original data. The physical parameters automatically listed are some of the standard items included in the international seismic bulletin (e.g., epicenter location, origin time, and magnitude). Results of this automatic method are presented.
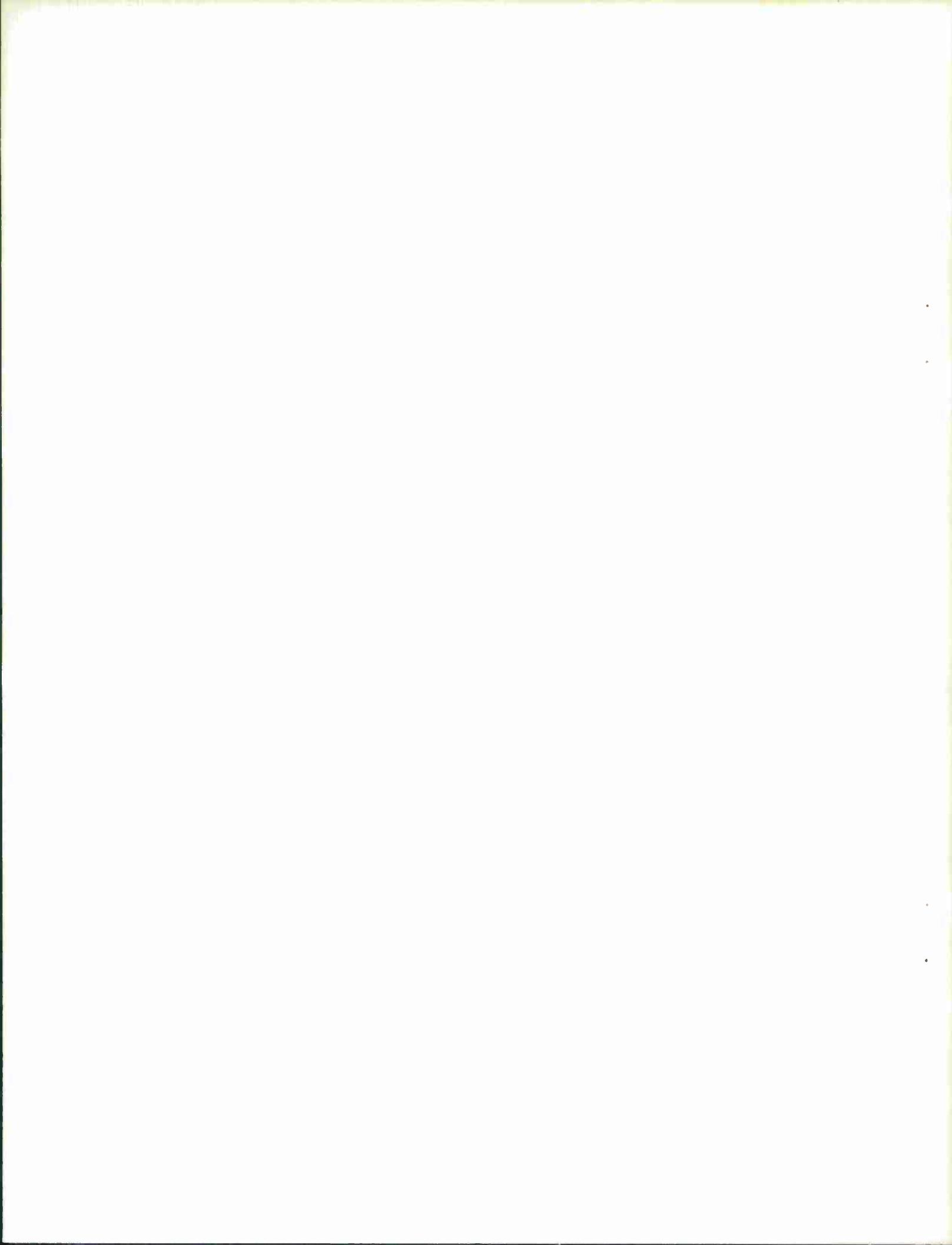
FASTABUL
(A Fast Automatic STAtion BULletin Program)

I.   INTRODUCTION

It is desirable to automate as much of the routine work at LASA as possible in order to reduce the number and training of the personnel required to staff a site, and to improve the results by doing a better job in the sense that there is no fatigue or inconsistency in the way parameters are measured (as is the case when humans are allowed as a link in the chain).   This report describes a computer program that has a standard LASA data tape as input, and outputs a "station bulletin."

The station bulletin computes and tabulates:

1.   The amplitude of the first arrival
2.   The period of the first arrival
3.   The complexity of the seismic waveform
4.   The direction of first motion
5.   The horizontal phase velocity
6.   The azimuth of the best fitting plane wave across the whole array.

From these last two quantities, the station bulletin also computes:

7.   The distance of the epicenter from LASA
8.   The latitude of the epicenter
9.   The longitude of the epicenter
10.   The arrival time of the first motion at LASA
11.   The origin time.

From items 1, 2, and 7 it computes:

12.   The magnitude of the event.

The basic problems in an automatic station bulletin are to have the computer recognize a P (or PkP) first arrival from the background noise, to measure accurately these arrival times at each subarray, and to do this as well as a human operator can. This program goes through a series of operations on the seismic waveforms and picks arrival times, culls these arrival times and discards the ones it thinks are bad, and

1
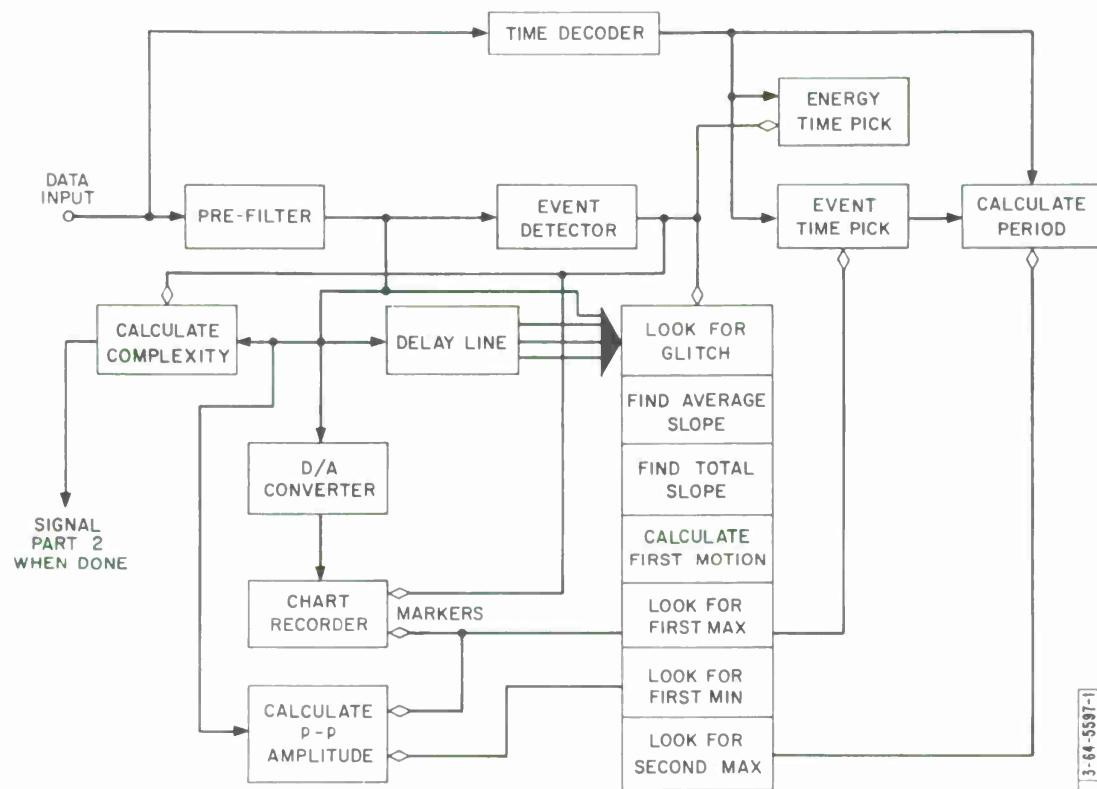
Fig. 1.   STABUL, Part 1.

computes the station bulletin. It also plots the seismograms and marks on them the time it picks. Under normal conditions, it does rather well, but sometimes, as when the signal is extremely weak, or when two teleseisms arrive together, it does poorly. In these cases, a human operator can look over the time picks and perhaps make a wiser choice for the computer to use to re-compute a new station bulletin based on these choices. This is called the "semiautomatic mode" of operation.

The input waveforms to this program can be the twenty-one "10" seismometer outputs, or 21 maximum-likelihood processed traces where each trace is a combination of the entire output from each subarray, or waveforms from any type of S/N improvement schemes.

Therefore, the minimum magnitude for which this automatic station bulletin will work depends on the type of preprocessing used on the input waveforms. If only one waveform is input to the program (for example, a beam made up from the entire array), the location of the source is determined by where the beam is pointed, so the program need only calculate complexity, magnitude, first motion, and origin time.

II. DESCRIPTION

The program is divided into three parts. Part 1 picks the event from the noise and calculates the amplitude, period, first motion, and complexity for each of the input seismograms. Part 2 decides which of the results computed by the first part are reliable, and Part 3 uses these reliable data to compute the station bulletin. Although this program is used off-line, the first two parts were designed to run in real-time on-line. This constraint means that the parameters must be calculated as the data come one sample at a time off the data tapes (or, in real-time, from the seismometers). This rules out such techniques as cross correlation and backing up the tapes to make repeated passes through the data.

Part 1 is iterated 21 times in the program, once for each subarray, each working by time sharing with the others.

Figure 1 shows a block diagram of the first part of the program. The data from the analog sum (chosen because it has better S/N than any individual seismometer) go into the standard energy event detector which has been described elsewhere.* When an

---

* H. W. Briscoe and P. L. Fleck, "A Real-Time Computing System for LASA," Proc. Spring Joint Computer Conf., Boston, April 1966.

3
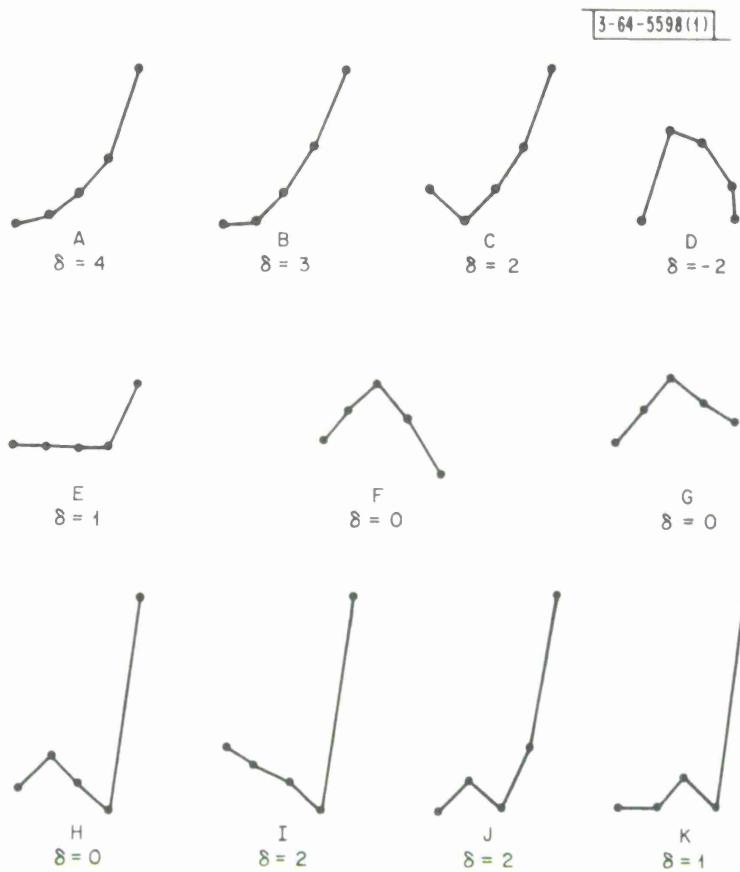
Fig. 2.   Some examples of positive first motion.

increase of energy in the proper frequency band (that of P or PkP phases − 1.2 to 1.7 cps) is present, the event detector triggers five operations.   First, it disconnects itself from having any further action on this channel.   Second, it decodes the GMT time from the tape and saves it as the "energy time pick."   Third, it starts the complexity calculation which runs by itself for the next 35 seconds.   Fourth, it puts a marker pulse on the analog chart recorder to mark this time pick.   Fifth, it starts the next phase of the program which picks the parameters from the waveform.

Since the first motion may have already occurred by the time the event detector triggers (especially when the event is weak), the preceding 200 milliseconds of data are saved in a delay line five samples long.   Some possible examples of what the program will find in the delay line at the event detector trigger time are shown in Fig. 2.

4

The program finds the average slope δ in the waveform by giving equal weighting to the four individual first differences between the five consecutive samples stored in the delay line. If any one of the first differences exceeds a threshold, the trigger was due to an impulse in the data (i.e., a false alarm), or if the channel is inoperative (all $\delta = 0$) (i.e., it has no energy), a "no good" flag is set which causes the data from this channel to be ignored in all the following calculations.

Next, the program calculates the over-all slope from the two data samples at either end of the delay line. If both slopes have the same sign, this sign is called the first motion. If the signs of the slopes differ, the "no good" flag is set for this channel. If there was a false alarm, the sign of the average slope is called the first motion. If there was no glitch, the over-all slope is called the first motion.

If the signal was weak, the situations shown in Figs. 2F or 2G can arise. The average slope is zero, and the over-all slope is equally likely to be up or down. In this case, the program uses the earliest individual slope as the first motion. If this slope is 0, then the program uses the over-all slope.

For all the cases shown in Figs. 2A to 2G, these procedures will yield the correct first motion. Cases in Figs. 2H through 2K will fail but these are unlikely to occur, since the data are band-limited and cannot change abruptly.

Knowing the first motion, the program examines the data, sample by sample as it comes from the delay line, and looks for the first maximum (or minimum). When it finds this, the program does three things:

      (a) Decodes the time and saves it as the "event time pick,"

      (b) Marks this time on the chart recorder,

      (c) Saves the amplitude of the waveform.

The program looks for the first minimum (or maximum). When it finds it, the program uses the amplitude of the waveform with (c) above to calculate the peak-to-peak amplitude. Finally, the program looks for the second maximum (or minimum). When it finds it, the time is again decoded and reduced by time of the first maximum to determine the period.

The peak amplitude, as calculated by this program, is defined as one-half of the difference of the signal amplitude between the first maximum and the first minimum after the event detector triggers. The period, as calculated by this program, is defined

5

1. FIND PREDOMINANT FIRST MOTION (PFM)

2. KEEP ONLY TIME PICKS CORRESPONDING TO PFM

3. DISCARD ALL PICKS FLAGGED BY THE BAD LIST

4. COMPARE DIFFERENCE BETWEEN MEDIAN PICK AND ALL PICKS
   WITH TABLE OF MAXIMUM ALLOWABLE DIFFERENCES

5. DISCARD PICKS WHICH FALL OUTSIDE THE MAXIMUM ALLOWABLE

6. USES ENERGY PICKS INSTEAD OF EVENT TIME PICKS
   IF THE AVERAGE AMPLITUDE IS 1 mμ OR LESS

Fig. 3.   STABUL, Part 2.

SUBARRAY "AO"
46° 41' 19" N
106° 13' 17" W
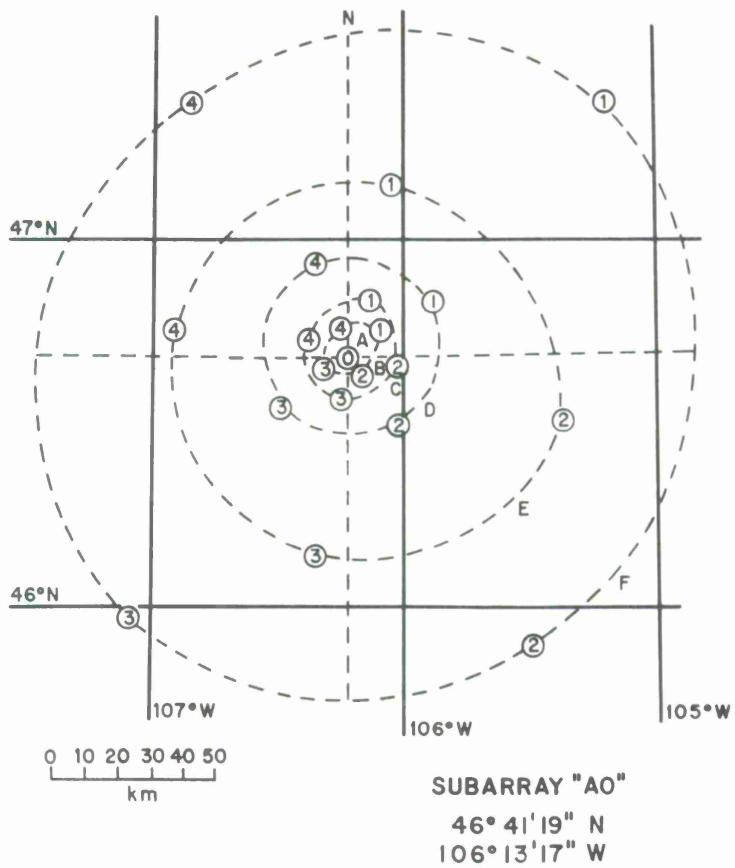
Fig. 4.   Map of LASA array.

6

as the time difference between the second and first maximum (or minimum, if the first motion is negative).

The purpose of the second part of the program is to examine the time picks and discard those which are unreliable. It is initiated when the complexity calculations are completed from all 21 sites. Since the complexity is defined as the inverse ratio of the rectified sum of the first 5 seconds of the event to the following 30 seconds of the event, this program will start 35 seconds after the last event has triggered the event detector. It does the operations listed in Fig. 3.

First, since the picks may be on different phases of the seismogram, only the set of picks for which the first motions are in the same direction are used.

Second, the picks from the channels which have dubious data and those for which the first motion calculation were in doubt are discarded, i.e., "no good" flag is set.

Third, since each seismogram is from a fixed geometric array pattern (Fig. 4), it is clear that the time picks from the B-ring must have less scatter than the picks from the F-ring. The program throws out picks which have a deviation from the median larger than could be caused by a genuine teleseism.

The program uses the good time picks to determine which channels to use in averaging the individual amplitudes, periods and complexities.

The third part of the program (Fig. 5) calculates the azimuth (β) and horizontal phase velocity ($\bar{v}$) of the plane wave which best fits (in the least-squares sense) the

1. CALCULATE $\bar{v}$ AND β FROM TIME PICKS

2. MODIFY THE TIME PICKS FROM STATION CORRECTIONS

3. RECALCULATE $\bar{v}$ AND β

4. CALCULATE Δ FROM $\bar{v}$

5. CALCULATE SOURCE LATITUDE AND LONGITUDE FROM Δ AND β

6. CALCULATE TRAVEL TIME FROM Δ

7. CALCULATE ORIGIN TIME

8. CALCULATE MAGNITUDE FROM A, τ, AND Δ

Fig. 5. STABUL, Part 3.

7

R - 0.2950 + (0.3418) SINE (BETA - (0.8727)) + (0.3209) SINE 2 (BETA - (2.8115))
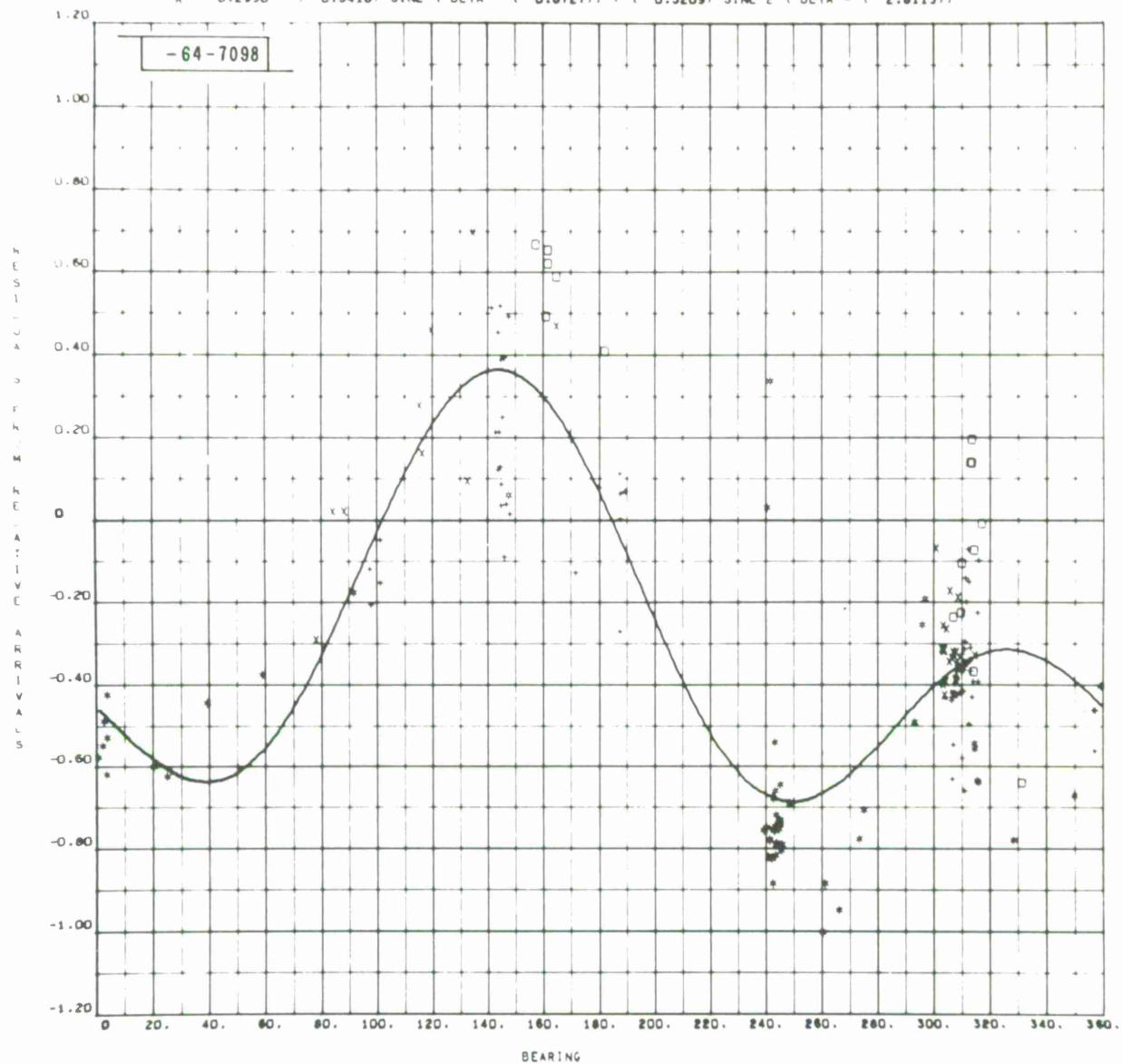

Fig. 6.   Station corrections for the pair F4–A0, plotted vs bearing for four intervals of distance.   The distance intervals are identified by the symbols as follows:

□ $0 \leqslant \Delta \leqslant 35°$ $+$ $55° \leqslant \Delta \leqslant 85°$

$\times$ $35° \leqslant \Delta \leqslant 55°$ $*$ $85° \leqslant \Delta \leqslant 105°$

8

selected time picks. The method used is described in detail elsewhere.* From this azimuth, station corrections are applied to the time picks. These station corrections are necessary because local inhomogeneities in the earth around the seismometer can cause variances in the velocity of the teleseism. These station corrections are mainly a function of azimuth and are approximated by a least-squares best-fitting Fourier series containing terms up to 2β (Fig. 6). The program then repeats the calculation of $\bar{v}$ and β using the modified time picks. One iteration is enough to determine $\bar{v}$ and β within the errors inherent in our present knowledge of the station corrections themselves.

In order to calculate the magnitude, origin time, and coordinates of the epicenter, one must use empirically derived curves of travel time vs distance, velocity vs distance, and "Q factor" vs distance. This program uses as a basis for these calculations the P and PkP travel time tables of Jeffreys-Bullen[†] for 33-km depth of focus. These raw data points have been doubly smoothed by finding the best-fitting quadratic to seven consecutive points and using the center point. A new point is generated by sliding the seven points down one point and finding a new quadratic, etc. This smoothing process was repeated by smoothing the smoothed data in an identical manner, and the derivative of the second iteration, i.e., the reciprocal velocity vs distance, was piecewise approximated by simple tangent, quadratic, and linear functions. The maximum error of distance vs velocity is less than $0.5°$ for distances between $11°$ and $90°$ and less than $2°$ for distances greater than $90°$ and less than $180°$. From these functions and $\bar{v}$, the distance between LASA and the epicenter is determined. It is then a simple problem in spherical trigonometry to calculate the epicenter latitude and longitude (see Kelly, op. cit., for details).

The magnitude is calculated by the standard equation:

$$m = \log_{10} \frac{A}{P} + Q(\Delta) \quad .$$

* E. J. Kelly, "Limited Network Processing of Seismic Signals," Group Report 1964-44, Lincoln Laboratory, M.I.T. (4 September 1964), DDC 447220.

† J-B curves from U.S. Coast and Geodetic Survey, 1966.
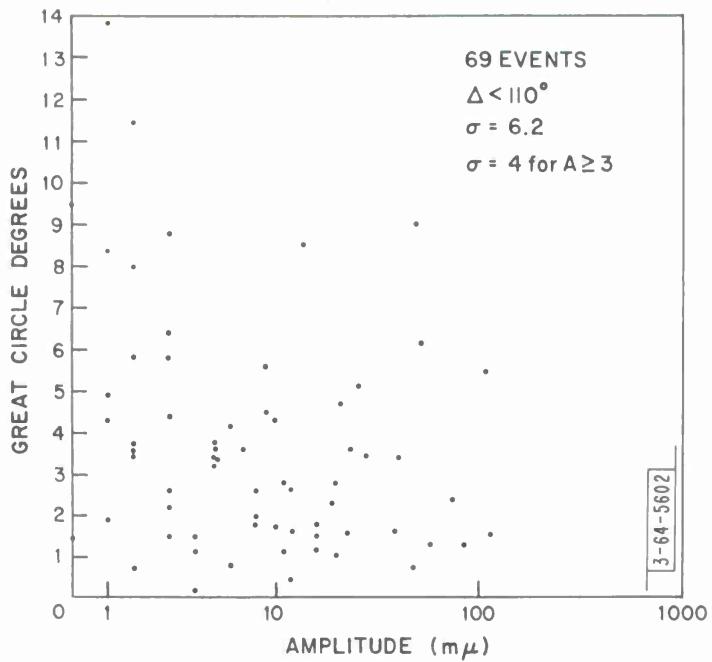
Fig. 7.   Error in location.
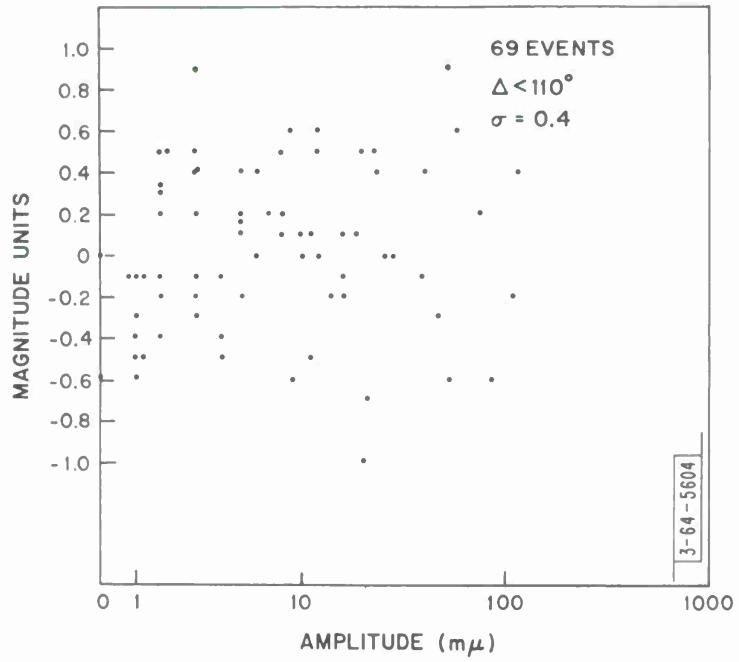
Fig. 8.   Error in magnitude.

10

The amplitude and period have already been determined, and the $Q(\Delta)$ is calculated from a piecewise linear approximation to the Jeffreys-Bullen Q factors for a depth of 33 km. The maximum error is less than 0.2 magnitude units for distances between 18° and 110°. The magnitude is not calculated for distances less than 18°. For PkP phases, the magnitude is approximated to about one magnitude unit.

The doubly smoothed travel time curves have been roughly approximated by piecewise quadratic functions for both P and PkP phases with a maximum error of 5 seconds. From these data, it is easy to calculate the origin time.

Finally, the program types out all the parameters it has calculated. If a human operator decides some of the automatic time picks are bad, he can start the program in the "semiautomatic" mode and manually tell the program which picks are bad. The program then uses Part 3 only to type out a new bulletin based on these picks.

III. RESULTS

This program has been tested in the fully automatic mode using 70 events which are all the events in our library from 1 October 1965 to 13 January 1966 that have been identified with CGS events and for which, therefore, we know the location, magnitude, and origin time. This library should represent a typical population of events. The events which were not saved in our library were later arriving phases of strong events.

The results of the test for all the events with epicenters less than 110° from LASA (P phases only) are shown in Figs. 7, 8, and 9. The error is plotted as a function of LASA amplitude because, as the signal becomes stronger, the automatic identification becomes better.

The over-all mean square error using fully automatic location is 6.2 great circle degrees. The mean square error for all the events which were received with an amplitude greater than 2 millimicrons is 4 degrees. These errors can be reduced somewhat by employing the semiautomatic mode, but there is a limit of about 2.5 degrees due to the limited aperture of the LASA array itself (200 km) and inherent timing errors which include our present imperfect knowledge of the station corrections. These timing errors become much more restrictive for the PkP phases. The errors due to the travel time curves seem small compared to these timing errors, even though a depth of 33 km is assumed.

11
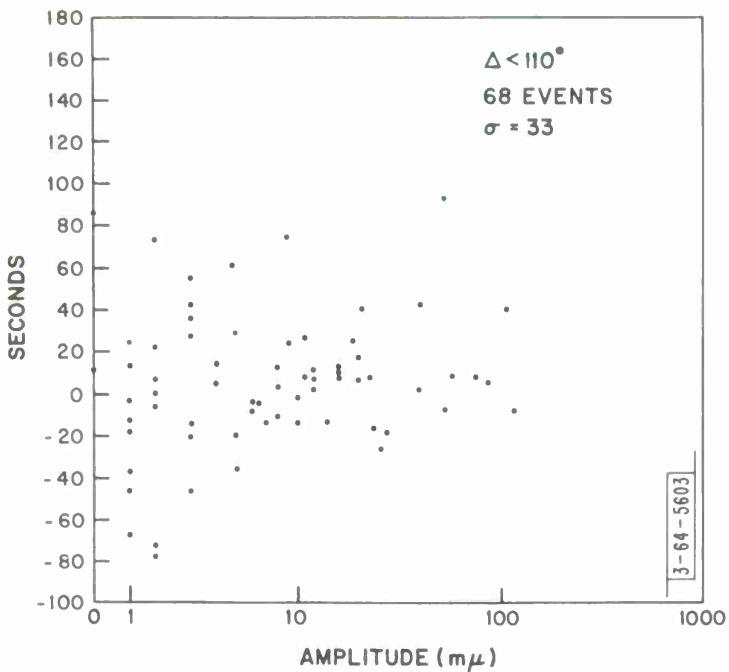
Fig. 9. Error in origin time.

The over-all mean square error in LASA magnitude is 0.4, but this is not surprising since the CGS reports are averaging stations with worldwide coverage and which have large deviations themselves. The over-all mean square error in determining the origin time is 33 seconds.

As mentioned earlier, it is difficult to put a lower bound on the magnitude for which this automatic station bulletin will work. If the signal is received with an amplitude greater than 5 millimicrons, the automatic picks are made about as well as a human can.

12

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
Lincoln Laboratory, M.I.T.	Unclassified
	2b. GROUP
	None

3. REPORT TITLE

FASTABUL (A Fast Automatic STAtion BULletin Program)

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Note

5. AUTHOR(S) *(Last name, first name, initial)*

Fleck, Philip L.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6 January 1967	18	3

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF 19(628)-5167	
b. PROJECT NO.	Technical Note 1967-3
ARPA Order 512	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	ESD-TR-67-16

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Advanced Research Projects Agency, Department of Defense

13. ABSTRACT

A real-time program for a digital computer is described which uses the sampled outputs of an array of 21 seismometers as input. These data are reduced and output as a typewritten list of the physical parameters of any "event" present in the original data. The physical parameters automatically listed are some of the standard items included in the international seismic bulletin (e.g., epicenter location, origin time and magnitude). Results of this automatic method are presented.

14. KEY WORDS

seismology	computer applications
LASA	real-time computer program